THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

The role of modality: Auditory and visual distractors in Stroop interference

Citation for published version:
Elliott, EM, Morey, C, Morey, RD, Eaves, S, Shelton, J & Lufti-Proctor, D 2014, 'The role of modality: Auditory and visual distractors in Stroop interference' Journal of Cognitive Psychology, vol 26, no. 1, pp. 15-26., 10.1080/20445911.2013.859133

Digital Object Identifier (DOI):
10.1080/20445911.2013.859133

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Author final version (often known as postprint)

Published In:
Journal of Cognitive Psychology

PLEASE CITE:

Elliott, E., Morey, C., Morey, R., Eaves, S., Shelton, J., & Lutfi-Proctor, D. (2014). The role of modality: Auditory and visual distractors in Stroop interference. *Journal of Cognitive Psychology*, *26*(1), 15–26. doi:http://dx.doi.org/10.1080/20445911.2013.859133

The Role of Modality: Auditory and Visual Distractors in Stroop Interference

Emily M. Elliott

Louisiana State University

Candice C. Morey, Richard D. Morey

University of Groningen

Sharon D. Eaves, Jill Talley Shelton, and Danielle A. Lutfi-Proctor

Louisiana State University

Author Note:

Emily M. Elliott, Department of Psychology, Louisiana State University; Candice C. Morey is currently in the School of Philosophy, Psychology, and Language Sciences at the University of Edinburgh; Richard D. Morey, Heymans Institute for Psychological Research, University of Groningen; Sharon D. Eaves is currently in the Department of Psychology at Stephen F. Austin State University; Jill Talley Shelton is currently in the Department of Psychology at the University of Tennessee at Chattanooga, and she was supported by the National Institute on Aging Grant T32 AG00030-32 during the preparation of this manuscript; Danielle A. Lutfi-Proctor, Department of Psychology, Louisiana State University .

We would like to thank Anh-Thu Vu and Leigh Grace Rouyer for their invaluable help with data collection.

Correspondence regarding this article should be addressed to: Emily Elliott, Department of Psychology, 236 Audubon Hall, Louisiana State University, Baton Rouge, LA, 70803, USA, phone (225) 578-7460, fax: (225) 578-4125, email: eelliott@lsu.edu.

Abstract

As a commonly used measure of selective attention, it is important to understand the factors contributing to interference in the Stroop task. The current research examined distracting stimuli in the auditory and visual modalities to determine if the use of auditory distractors would create additional interference, beyond what is typically observed in the print-based Stroop task. Research has supported additive effects of auditory and visual distractors; however, there is only one empirical demonstration of this finding to date. Using different versions of the Stroop color naming task, behavioral analyses of reaction times (RT) were conducted, along with distributional RT analyses. The results indicated that a combination of visual and auditory distraction did not lead to a larger interference effect than visually-based distraction alone. These findings suggest that methodological issues may have influenced the prior finding of additive effects of the two modalities, and are discussed in relation to the word production architecture account of Stroop effects.

The Role of Modality: Auditory and Visual Distractors in Stroop Interference

Choosing which stimulus to attend to in a multisensory array is a daily problem that many individuals navigate successfully. However, there are times at which selective attention processes fail. Failures of selective attention have been investigated in both laboratory and applied settings for many years, and the Stroop task is a popular choice for such investigations. As an example, selective attention is required for successful performance on incongruent trials in a typical, visual Stroop task; if the word "blue" is presented in red ink, the correct color naming response would be "red" (Stroop, 1935). Participants are reliably slower to name the ink color of a conflicting color word than a neutral or congruent color word. This pattern of slowed responding has been termed *Stroop interference*.

Researchers have also investigated Stroop interference within the auditory modality. Green and Barber (1981) found that participants were slower to judge the gender of a speaker if a female speaker said the word "man" than if the female speaker said "girl". The authors suggested commonalities between the auditory and visual versions of the Stroop task, a finding supported by neuroscientific research as well (Donohue, Liotti, Perez, & Woldorff, 2012; Roberts & Hall, 2008). The underlying mechanisms of Stroop interference therefore appear to be similar, regardless of modality. Both auditory and visual Stroop effects are linked to a semantic mechanism and a response competition mechanism, as revealed, for example, by semantic gradient effects in both modalities and interference effects when stimuli are incongruent (Green & Barber, 1981; Risko, Schmidt, & Besner, 2006; Schmidt & Cheesman, 2005).

Further, both auditory and visual Stroop tasks have been applied in many domains of the psychological literature as an index of selective attention processes (Donohue et al., 2012; MacLeod, 1991), and different versions of the task allow for different research questions to be addressed. In addition to these single modality varieties of the Stroop task, a cross-modal version has also been investigated which incorporates a visual target with auditory distractors (Cowan & Barron, 1987). However, the focus on the underlying mechanisms of cross-modal Stroop has received much less attention in the literature. The goal of the current study is to examine cross-modal Stroop effects to determine the role of the modality of the distracting stimuli. If the cross-modal Stroop task shares properties with the single modality versions, one would expect similar patterns of findings to emerge from the cross-modal version as from single modality versions.

Cross-Modal Stroop

The cross-modal variant of the Stroop task includes distractors in the auditory modality, as opposed to the traditional, printed Stroop task in which both the target and distractor are visually presented, in one stimulus item. Although the interference effects are typically smaller than those observed in the traditional Stroop paradigm, the effects are reliable and have been replicated in adults and children (Cowan & Barron, 1987; Elliott, Barrilleaux, & Cowan, 2006; Elliott & Cowan, 2001; Elliott, Cowan, & Valle-Inclan, 1998; Hanauer & Brooks, 2003; Roelofs, 2005; 2012; Shimada, 1990). The cross-modal Stroop task was first investigated by Cowan and Barron, and the visual stimuli presented for color naming were either neutral (i.e., rows of x's), or incongruent (i.e., color words printed in differing colors of ink). Participants were asked to name the colors of these stimuli as quickly as possible, while ignoring auditory distractors. These irrelevant auditory stimuli included color words chosen from the same response set as the targets. The results indicated significant impairment in color naming in the presence of auditory color words for both types of visual stimuli; this finding was interpreted as an effect of generalized auditory distraction because it occurred for both the incongruent and neutral visual stimulus types.

Furthermore, performance was the slowest and the most error-prone when participants named the ink color of incongruent color words while hearing irrelevant color words. These results were interpreted to suggest an additive model of Stroop effects in the cross-modal version of the task. Stroop interference similar to the traditional, print-based Stroop task was observed (i.e., color naming was interfered with by written color words), generalized auditory distraction was observed, and the presence of stimuli in the two modalities increased the interference effect over and above either type of interference alone. The cause of the additive interference was ascribed to the utilization of a buffer in memory, which held items from both the auditory and visual modalities, in the service of preparing a spoken response to the color naming task. With distractors in both the visual and auditory modalities, the task of response selection was made more difficult than it would have been with a distractor in only one modality, as there were more items to search through when preparing a response (Cowan & Barron, 1987).

An interesting series of replication attempts and rebuttals followed the publication of the original cross-modal Stroop research (Cowan, 1989a; Cowan 1989b), and questioned the existence of the buffer (Miles & Jones, 1989; Miles, Madden, & Jones, 1989). However, later research using the cross-modal Stroop task introduced methodological changes to allow for precise control of the onsets of the auditory and visual stimuli, and determined that the stimulus timing was important to the effect (Elliott et al., 1998; Shimada, 1990). Furthermore, the visual stimuli were modified by presenting colored squares for naming (Elliott et al., 1998), making the task appropriate for participants who cannot read (i.e., children; Hanauer & Brooks, 2003).

Consistent with other research using the cross-modal Stroop task, the findings from Elliott et al. (1998) indicated a clear role for both interference effects and auditory distraction; Indicating interference, responses were significantly slower in the incongruent color word condition than in the non-color word condition. Furthermore, indicating generalized auditory distraction, the non-color word condition was significantly slower than the silent condition (i.e., the pattern of mean RT's indicated silent < non-color < color). These effects were shown only when the onsets of the two types of stimuli were simultaneous; when the auditory stimulus preceded the visual by 500 ms, there were no interference effects of either kind. Instead both the incongruent color word and non-color word auditory conditions were faster than silence (i.e., color = non-color < silence), and the authors suggested that participants used the auditory stimuli as a cue to the visual target's onset. However, the Elliott et al. experiment used colored squares as targets, as opposed to print-based, colored visual stimuli. Thus, it was not possible to evaluate Cowan and Barron's (1987) earlier finding of additive interference effects from the visual and auditory distractors.

Cross-Modal Stroop: One Distractor or Two?

The motivation for Cowan and Barron's (1987) modification to the original Stroop paradigm was centered more on an understanding of the structure of the working memory system than on the basis of Stroop effects themselves. However, one could conclude from their findings that two mechanisms of interference were present in the cross-modal version of the task: a form of response competition on the one hand, and auditory distraction on the other. The finding of two mechanisms contrasts with the view of Roelofs (2005). The word production architecture account (Roelofs, 2003; 2005) has been applied to both the traditional, printed Stroop task and the cross-modal version. It draws upon the differences in the processes involved when naming a color as compared to naming a word. This account can explain the asymmetry of RTs when participants perform these two tasks: word reading is performed with no intermediate processing steps because the item is already in a form that is suitable for a verbal response, whereas color naming must be performed by converting the colored item into a label that can be spoken. This difference in processing can explain why participants are slower to name the color of ink when the printed word is incongruent with the ink color, but are not slower to read the written word regardless of the ink color. The pattern of results has been referred to as the “color-word Stroop asymmetry” (Roelofs, 2005, p. 1325).

Roelofs’ research suggested that a similar asymmetry could be found in the cross-modal Stroop task, and the finding of the asymmetry in both the traditional and cross-modal versions of the task would support a common mechanism of Stroop effects. In a series of experiments, spoken word naming was not slowed by a visually-presented colored square, but naming a visually-presented colored square was slowed by an incongruent auditory color word. He argued, based on these findings, that differences in the functional architecture of color naming and spoken or written word naming are driving performance in both the traditional Stroop task and the cross-modal version (Roelofs, 2005). Within this interpretation, the modality of the distractor in the Stroop task is not the underlying cause of the interference effects observed, and there was no specific discussion of generalized auditory interference effects.

Roelofs’ experiments (2005) contrasting spoken word and written word distractors have provided key information about the interference observed, and have suggested that the color-word Stroop asymmetry applies to both the traditional, printed version of Stroop interference, as well as the cross-modal version of Stroop interference. However, the conditions from Cowan and Barron’s (1987) original experiment have not been replicated exactly; there has been no further empirical demonstration of printed words as distractors with spoken auditory distractors when the task was color naming of the printed word. Without a direct assessment of auditory and visual distractors within one Stroop task, the question of whether additive effects of interference from the use of both auditory and visual distractors occur cannot be answered.

The Current Study

The current study included four versions of the Stroop task to investigate the role of the modality of the distractors, as well as the type of visual stimulus used for color naming (see Figure 1). A traditional, printed version of the Stroop task was included for a comparison to previous literature. Additionally, three versions of the cross-modal Stroop task were investigated with different types of visual stimuli: colored squares, colored @ symbols, and printed items. The version with the colored squares replicates prior research with cross-modal Stroop tasks (Elliott et al., 1998; Roelofs, 2005), and the version with colored @ symbols extends this prior research to a type of nameable stimulus that is not a printed word. For ease of reference, the version with the printed items was termed “multi-modal”. It most closely matched the previous work by Cowan and Barron (1987); in the current design the printed targets were colored words accompanied by auditory distractors that matched the visual distractor dimension, and participants were asked to name the color of ink as quickly and accurately as possible. These three versions of Stroop tasks with auditory distractors allowed a comparison of semantics (through the type of visual target), response competition (through the inclusion of congruent and incongruent trial types), and generalized auditory distraction (through the use of word and non-word auditory stimuli). Finally, all Stroop versions included a high proportion of congruent trials to maximize the size of the interference effects (Logan & Zbrodoff, 1979; Meier & Kane, 2013).

Figure 1 about here

Distributional analyses were conducted in the current study, in addition to mean analyses of reaction times, to obtain a more nuanced understanding of response times across the different versions of the Stroop task. We chose a graphical technique called the delta plot (De Jong, Liang, & Lauber, 1994; Pratte, Rouder, Morey, & Feng, 2010; Ridderinkof, Scheres, Oosterlaan, & Sergeant, 2005; Speckman, Rouder, Morey, & Pratte, 2008). Delta plots are built using the quantiles of the RT distribution, and previous research has revealed, that even without a precise statistical inference test, delta plots provide key information about the underlying mechanisms of tasks thought to assess constructs such as cognitive control and inhibition (Pratte et al., 2010; Unsworth, Redick, Spillers, & Brewer, 2012). With delta plots, we compared the interference effects across the different versions of the Stroop task, to determine if similar patterns would be obtained, regardless of version, thus ensuring that the statistical inferences we made on averaged responses did not obscure conflicting evidence that could be discerned from the response time distributions. For example, observing a positive slope in which the size of the interference effect increases as response times increased in all four tasks would be consistent with the hypothesis that there is a single underlying mechanism driving the interference effect. Findings from the delta plots lend further support to the inferential claims made based on analyses of averaged response times.

Method

Participants

Two hundred Louisiana State University undergraduate students participated (age: M = 19.69 years, SD = 1.80) for course credit or extra credit in psychology courses. Individuals were not eligible to participate if they reported abnormal hearing or vision, use of medications that alter cognition, a first language other than English, or an outlying age (a 41-year old was excluded). Data from one participant was excluded for failure to follow instructions.

Materials and Design

Typical Stroop conditions were included in the design of the current research (e.g., both congruent and incongruent color conditions) as well as a noncolor word condition and a silent or "neutral" condition, depending on whether the version of the task was unimodal or cross-modal.

The experiment utilized a between-subjects design, with each participant randomly assigned to one of three of the Stroop tasks: traditional Stroop (n = 47), cross-modal Stroop (n = 49), or multi-modal Stroop (n = 48). Data for cross-modal @ Stroop (n= 56) were collected separately at a later time, with a new group of participants sampled from the same population. Within each task, the following distractor conditions were used: congruent (color of object matched the color word presented visually and/or aurally), incongruent (color of the object did not match the color word presented), non-color (word presented is not a color word), and neutral/silence (no auditory distractor and/or visual distractor is present).

All tasks were presented using E-Prime software (Schneider, Eschman, & Zuccolotto, 2002) on Dell Dimension desktop computers with 17-inch monitors. For all Stroop tasks, the participants' RTs were recorded by a headset microphone connected to a response box that logged the vocalization onsets. The auditory distractors used in the cross-modal and multi-modal Stroop tasks were presented through headphones as a digitized female voice, which lasted 210-500 ms, and were measured with Quest sound-level meter and earphone coupler in the range of 77-81 dB(A). The task was completed in a 20-30 minute session in a room with only the participant and experimenter.

Traditional Stroop task. This task was a computerized version of the original Stroop task, and no auditory distractors were presented. The participant was instructed to name the color of the word presented on the screen, and to ignore all other information. The target stimuli colors were: *red*, *blue*, and *green*. The distractors corresponded to the four trial conditions. In the congruent trials, the displayed color matched the color word presented (e.g. "red" presented in red). In the incongruent trials, a color word was presented that did not match the displayed color (e.g. "red" presented in green). For the noncolor trials, the visual stimulus was taken from the category of size words (*big*, *long*, or *short*) and for the neutral/silence condition, four @ symbols were used to present the color.

Cross-modal Stroop task. This task was similar to the Stroop task discussed above, except that the cross-modal task used aural distractors instead of visually-presented words. Participants were instructed to ignore anything heard in the headphones and to name the color of the 4.4 cm x 4.4 cm square. The auditory distractors were presented simultaneously with the square and again corresponded to the four conditions: congruent, (where the color word heard matched the color of the square), incongruent (the color word heard did not match the color seen), noncolor (the word heard was one of the designated non-color words), and neutral/silence (the square was presented with no auditory stimulus).

Cross-modal @ Stroop task. This task was identical to the cross-modal variant described above, with the exception that the colored squares were replaced with colored strings of @@@@. These were chosen as nameable yet non-word-like visual stimuli, and to facilitate baseline comparisons with the neutral conditions of the Traditional and Multi-modal Stroop tasks, which also presented colored @ symbols.

Multi-modal Stroop. This task combined the traditional Stroop task with auditory distraction. The instructions were to name the displayed color of the printed word and to ignore anything heard through the headphones. For example, in the congruent condition, the word "red" would be presented in red, as the participant heard the word "red" through the headphones. The auditory distractor condition always matched the visual distractor condition. In the incongruent condition, participants were asked to name the color of a printed word while ignoring both the word itself and the auditory distractor (e.g., the word *red* displayed in blue with the spoken word "red" should lead to the response of "blue"). The noncolor condition presented the printed noncolor word in color, and was accompanied by the same spoken noncolor word as well. Finally, in the silent/neutral condition no auditory distractors were presented, and participants named the color of @@@@.

Procedure

After the participant received instructions, 24 trials were presented for practice using the microphone and naming the visual stimuli, with no distractors. The experimental portion of the task consisted of two blocks of 135 randomly ordered trials with a short break in between. Following the convention of previous research with a manipulation of congruency percentages (Kane & Engle, 2003; Meier & Kane, 2013; Morey et al., 2012), unanalyzed filler trials were included to satisfy the 75% overall congruency proportion within the task. These unanalyzed filler trials were labeled only within the Eprime program, and were not identified during the presentation of the experimental stimuli as such. There were 99 congruent trials in each block, divided into 12 non-filler trials and 87 filler trials. There were 12 incongruent trials in which each of the possible 6 incongruent combinations was used twice, 12 non-color trials in which each of the following 6 noncolor combinations was used twice: *red-big*, *red-long*, *blue-long*, *blue-short*, *green-big*, *green-short,* and 12 neutral/silence trials.

Each trial began with a fixation cross which remained on the screen for 500 ms, then the target was presented on the screen and remained until the microphone detected a response. The participant was asked to name the color of the object presented as quickly and accurately as possible. The experimenter used the keyboard to respond to three questions following each trial. They were asked to record the color word said by the participant, indicate a false start by the participant (triggering the microphone with an incomplete response), and indicate whether any errors were made by the experimenter in answering the previous two questions.

Results

The results are divided into the analysis of the error rates, followed by RT analyses using ANOVAs, and RT distribution analyses. The basic analyses each began with a 4 (between subjects: Stroop type: cross-modal, cross-modal @, multi-modal, and traditional) x 4 (within subjects: distractor condition: congruent, incongruent, silence, and noncolor) mixed ANOVA and were then followed by one-way ANOVAs, in cases of significant interactions. The *F* values are conventionally significant at the $p < 0.05$ level, unless otherwise reported. All of the analyses reported failed to meet the assumption of sphericity as calculated with Mauchly's W coefficient, so the Greenhouse-Geisser correction was used. All significant main effects were followed by pair-wise comparisons with a Bonferroni correction.

Table 1 about here

Error Analyses

The average experimenter error rate, errors that occurred when the experimenter pressed an incorrect key immediately after the participants' response, was less than 1% of trials across all Stroop tasks. The means for participants' response errors, by distractor condition and Stroop version, are presented in Table 1. The false start error rate, errors that occurred due to a pre-emptive sound from the participant before an answer was given, averaged about 3% of trials across the different Stroop versions. Further analyses on false starts and experimenter errors were not included.

A 4 x 4 mixed ANOVA was then performed on the participants' response errors, after the experimenter errors and false start errors were removed from the data set, across the different Stroop versions. There was a significant main effect of Stroop type, F (3, 196) = 12.55, MSE = 0.01, η_p^2 = 0.16, and the follow-up Bonferroni corrected pair-wise comparison indicated that the most errors were in the traditional version (M = 4.36%). The multi-modal (M = 2.89%) and cross-modal (M = 1.77%) versions did not differ significantly from each other (p = .22). Finally, the cross-modal @ version (M = 1.34%) differed significantly from the traditional and multi-modal versions, but not the cross-modal version (t < 1). There was also a significant main effect of distractor condition, F (1.34, 262.38) = 92.66, MSE = 0.01, η_p^2 = 0.32. These main effects were qualified by a significant interaction, F (4.02, 262.38) = 22.31, MSE = 0.01, η_p^2 = 0.26.

This interaction was examined with separate ANOVAs for each Stroop type. The cross-modal version had a significant main effect of distractor condition, F (2.04, 97.99) = 4.31, MSE = 0.01, η_p^2 = 0.08. Although no significant differences were found in the pair-wise comparisons, the means suggested that the largest errors occurred in the incongruent condition. The analyses for the cross-modal @ version were not significant, while the multi-modal version had a significant main effect of distractor condition, F (1.21, 56.90) = 45.56, MSE = 0.01, η_p^2 = 0.49;

indicating that the incongruent condition led to the highest error rate while the other conditions did not differ from each other (all *t*s < 1). The traditional version also had a main effect of distractor condition, *F* (1.10, 50.44) = 43.75, *MSE* = 0.02, $\eta_p^2 = 0.49$, and indicated the same pattern as the multi-modal version, with the most errors occurring during incongruent trials. The Bonferroni corrected comparison of errors in the non-color and congruent condition was nonsignificant ($t_{(138)} = 2.20$, $p = .12$) as well as the silence/neutral and non-color comparison ($t_{(138)} = 3.00$, $p = .05$) and the congruent and silence/neutral comparison ($t < 1$).

Figure 2 about here

RT Analyses

All incorrect responses were removed from the data prior to analyzing the RTs. Means of the medians were used to avoid the problems with skew typically associated with RT data. A 4 x 4 mixed ANOVA was conducted to determine how the versions of the Stroop task differed in overall RTs and in the magnitude of the interference effects produced (see Figure 2). The analysis revealed a main effect of Stroop type, *F* (3, 196) = 37.37, *MSE* = 26185.05, $\eta_p^2 = 0.36$. The cross-modal versions (squares *M* = 512 ms; @@@@ *M* = 544 ms) did not differ (*p* = .26) and had significantly faster RTs than the multi-modal (*M* = 651 ms) and traditional (*M* = 645 ms) versions, which also did not differ ($t < 1$). There was also a main effect of distractor condition, *F* (2.49, 487.47) = 466.95, *MSE* = 2364.90, $\eta_p^2 = 0.70$. All of the distractor conditions were significantly different from one another, congruent < silent < noncolor < incongruent. The main effects were qualified by a significant interaction, F (7.46, 487.47) = 39.52, MSE = 2364.90, η_p^2 = 0.38.

The task by distractor condition interaction was investigated with 4 one-way ANOVAs for each Stroop version. The cross-modal squares version had a significant main effect of distractor condition, F (2.02, 96.97) = 89.18, MSE = 2336.55, η_p^2 = 0.65. The main effect can be represented by the following: congruent = silent < incongruent < noncolor. Interestingly, in this version of the Stroop task the largest interference came from the noncolor condition, and the congruent and silent conditions did not differ significantly (t < 1). Given the large proportion of color word trials, relative to the noncolor word trials, we investigated the four distractor conditions by trial blocks, to determine if the noncolor words were acting as an “oddball” or deviant sound (Hughes, Vachon, & Jones, 1997; Morey et al., 2012). Such sounds tend to cause a neurological and/or behavioral response, due to their unexpected nature (see Parmentier, Elsley, Andrés, & Barceló, 2011; Schröger & Wolff, 1998). The 2 x 4 repeated-measures ANOVA revealed significant main effects of both trial block, F (1, 48) = 42.75, MSE = 4355.49, η_p^2 = 0.47, and distractor condition, F (2.08, 100.01) = 111.25, MSE = 4189.97, η_p^2 = 0.70, as well as an interaction of these two factors, F (2.66, 127.62) = 4.89, MSE = 1234.48, η_p^2 = 0.09. The interaction is depicted in Figure 3, with a larger decrease in RTs from trial block 1 to 2 in the noncolor word condition than in the other distractor conditions.

Figure 3 about here

Returning to the one-way ANOVAs for each Stroop version, while the cross-modal @ task also had a main effect of distractor condition, F (2.72, 149.85) = 79.04, MSE = 1510.62, η_p^2 = 0.59, the pattern of effects differed from the cross-modal squares version. The main effect can be represented by the following: congruent = silent < incongruent = noncolor. Because the noncolor word condition was not significantly slower than the incongruent condition ($t < 1$), the additional block analyses were not conducted. As in the cross-modal squares version, the congruent and silent conditions did not differ significantly ($t_{(165)} = 1.71$, $p = .55$).

The multi-modal version had a significant main effect of distractor condition as well, F (1.75, 82.03) = 112.12, MSE = 5906.67, η_p^2 = 0.71. This main effect differed from the cross-modal version's RT main effect in that the incongruent condition led to the slowest RTs (congruent = silent < noncolor < incongruent). As with the other two versions with auditory distractors, the congruent and silent conditions did not differ significantly ($t < 1$). Finally, the traditional version also had a significant main effect of distractor condition, F (2.18, 100.19) = 224.30, MSE = 4294.01, η_p^2 = 0.83. This main effect indicated that all conditions differed significantly, in the order of congruent < silent < noncolor < incongruent, and also revealed the only significant facilitation effect among the versions of the Stroop task.

To explore the observed differences in mean RT further, we created delta plots based on the averaged quantiles for each distractor type and task. It is typical for Stroop effects (a quantile-by-quantile comparison of the congruent and incongruent distributions) to increase in magnitude as response times increase. In a delta plot, this comparison appears as a positively-sloped line (e.g., Pratte et al., 2010). First, we examined whether comparing the incongruent and congruent distributions showed this characteristic pattern for each of our tasks, which we consider to be a necessary condition for arguing that the consistent slowing we observed with incongruent stimuli across Stroop tasks results from a similar cause.

Figure 4 about here

See Panel A of Figure 3 for the delta plots for incongruent-congruent distributions in the two cross-modal tasks, along with the multi-modal, and traditional Stroop tasks. Each of these tasks shows the positive scale change, with the difference between incongruent and congruent trials increasing throughout the RT distribution. This visual depiction makes clear that this pattern is stronger and effect sizes larger for the traditional and multi-modal versions of the task than for the cross-modal versions, though note also that the sizes of the Stroop effect in the cross-modal tasks reached approximately 100 ms. These plots further show that the range of onset response times differs, such that the earliest cross-modal responses are faster than the earliest responses in the other two tasks.

In the remaining panels of Figure 3, we compared incongruent, non-color word, and congruent distractor types against neutral or silent trials in each task, to explore the differences among the distractors uncovered by the ANOVAs further. Panel B shows the incongruent versus neutral comparison for each task. The characteristic increase of the effect of incongruency as RT increases is clear for each trace, especially for the multimodal or traditional versions. In Panel C, noncolor distractors are compared with neutral or silent conditions. This comparison provides information about the nature of interference from general auditory distractor effects and "generic" interference in the traditional version, which did not contain any auditory distractors.

The slopes of the four tasks are very similar here; much more so than in the previous two panels, suggesting that all of the task versions were subject to interference from the noncolor words. The final panel presents the comparison between congruent and silence/neutral trials, indicating the lack of general auditory distraction or facilitation in the cross-modal tasks. Consistently with the mean RT inferential analyses, throughout the distributions, hearing a congruent color word provokes no interference compared with silence, suggesting that the interference observed in the cross-modal task in the incongruent and noncolor word conditions cannot simply reflect an effect of hearing irrelevant speech (see Elliott & Cowan, 2005 for a discussion of irrelevant speech effects). Alternatively, in the traditional version of the task, a constant facilitation effect is apparent, while in the multimodal task, congruency might provide facilitation in the fastest trials, but this facilitation gradually disappears as RTs increase.

Discussion

The main question motivating the current research centered on the modality of the distractors in the Stroop task, and whether the use of two types of distractors would increase the magnitude of Stroop interference. Comparing four versions of Stroop-like color-naming tasks, we observed increased RTs in incongruent conditions relative to silence/neutral conditions when the source of interference was print-based (traditional Stroop), when the source of interference was auditory (cross-modal Stroop, cross-modal Stroop @), and when the source of interference included both auditory and print sources of interference (multi-modal Stroop). We also found that adding auditory distractors to the traditional print-based Stroop task did not produce greater interference from the two sources of distractors (e.g., written and spoken words), relative to only one distractor. Across tasks, we found that both spoken and written words interfered with color naming; a finding consistent with Roelofs' word production architecture account (2005).

To place the main findings of the current study in the broader context of previous Stroop research, it is important to revisit a few details. The traditional Stroop version of the task produced a clear interference effect on RT (i.e., the comparison of incongruent and congruent trials), as well as a clear facilitation effect (i.e., the comparison of neutral and congruent trials). No other version of the task produced a significant facilitation effect. The multi-modal version of the task presented a unique combination of visual and auditory distractors presented as matching written and spoken words; a task condition not tested since the original demonstration of the cross-modal Stroop task by Cowan and Barron (1987). As mentioned above, significant interference was noted, but the magnitude of the interference was not larger than what was observed in the traditional version of the task.

One candidate to explain this finding of reduced distractor interference in the multimodal version comes from research on the Stroop dilution effect, in which the presence of an additional neutral stimulus in the visual display reduces the size of both the interference and facilitation effects (Kahneman & Chajczyk, 1983). The original work on the Stroop dilution effect was restricted to visual stimuli, but researchers have investigated an auditory version of the Stroop task, and have found evidence of a dilution effect as well (Dittrich & Stahl, 2011). If one views the current multi-modal version of the task as a traditional, printed task with additional auditory distractors, then the concept of dilution is clearly evident: dilution of the interference effect occurs with the presence of additional distractors.

The current work extends previous investigations of the dilution effect with auditory stimuli to a multi-modal form, with visually-presented targets and both auditory and visual distractors. The effect of dilution may include decreased facilitation in the cross-modal task as well, but this hypothesis warrants empirical investigation as dilution itself has not been investigated in a cross-modal Stroop paradigm. However, further evidence for the dilution account comes from the finding that the mean RTs of the traditional and multi-modal versions did not differ, and thus separates this interpretation from a perceptual load-based account (e.g., Lavie, 2005), in which the added distractors reduce the attentional resources available for distractor processing. A final point to consider, with regard to facilitation effects, is that the current research focused on simultaneous presentation of the targets and distractors. Previous research has indicated that the time-course of facilitation effects may vary (see Roelofs, 2010), and it is an important direction for future research to examine the methodological conditions of the current study under a wider range of presentation timings for the targets and distractors.

The issue of presentation timing may also be relevant to the difference between the current findings of the multi-modal version and the original work of Cowan and Barron (1987). The timing of the auditory and visual stimuli is a major factor in the size of the interference effect; Elliott et al. (1998) found no interference effects when the auditory stimuli preceded the visual by 500 ms, but significant interference effects when the two types of stimuli had simultaneous onsets (see also Roelofs, 2005, for a discussion of the time-course of cross-modal Stroop interference). In addition, the measurement of the RTs and error rates were based on individual stimuli in the current research, as opposed to the sets of 100 items that were used in Cowan and Barron’s design. Finally, because the current design allowed for individual stimuli to be presented, it was the case that the auditory and visual distractors were always matched (spoken word and written word, “red”, written in blue font). It is possible that unmatched combinations would have emerged in Cowan and Barron’s study, such that the spoken auditory distractor did not match the written word distractor, leading to greater interference when both spoken and written distractors occurred, relative to either type of distractor alone. However, in terms of a dilution account, it is not clear if the matching or mismatching of the distractors would clearly influence the outcome. The timing of the stimuli may have been the largest factor differentiating the work of Cowan and Barron from the current study.

Moving now to the remaining cross-modal versions of the Stroop task, squares or @ symbols were used as the color-naming targets, but RTs did not differ significantly between these two versions. Although the visual target of @ symbols was clearly nameable, this factor did not produce significant change compared to RTs to colored squares as the visual target. Overall, faster RTs and smaller interference effects were observed in these two versions relative to the two versions with printed word distractors. One explanation for this pattern of results is the spatial separation of the targets and distractors in the cross-modal versions, which is known to decrease the size of the interference effect (Roelofs, 2012; Spieler, Balota, & Faust, 2000), as opposed to the spatial integration of the visual targets and distractors used in the multi-modal and traditional versions of the current experiment. Future research should manipulate the degree of spatial separation versus integration in a cross-modal version of the Stroop task to test this hypothesis directly.

Additionally, the use of @ symbols as the naming stimulus created the opportunity for a baseline comparison to the multi-modal and traditional versions in the current research. In the neutral/silent conditions across these three versions of the task, participants were asked to name the color of the @ symbols, in silence. Thus one might have expected comparable RTs across the three versions. However, inspection of Figure 2 reveals that the mean RT in the neutral/silent condition was faster in the cross-modal @ version than in either the multi-modal or traditional versions. Having the @ symbol as the only visual stimulus for color naming led to faster responses than having a combination of printed words and @ symbols as visual targets. This baseline comparison extends the effects of task context that have been found in previous Stroop research, in which manipulations of experimental parameters can influence patterns of responding (e.g., such as list-wide proportions of congruency, Hutchison, 2011; Meier & Kane, 2013; but see Schmidt & Besner, 2008, for an alternate account based on contingency learning).

Conclusions

Four versions of the Stroop task were contrasted, and consistent, robust slowing was observed with semantically incongruent distractors, regardless of distractor modality. The inclusion of auditory stimuli in the multimodal version reduced the interference effect typically observed with printed stimuli, relative to the traditional version, and was interpreted with respect to previous findings of Stroop dilution effects. Other factors, such as the degree of spatial separation versus integration, and the dilution effects mentioned above, contributed to some of the observed differences in the task versions. Continued use of multiple versions of the Stroop task seems appropriate, given that some of the characteristics of the different versions may be better suited for certain populations over others.

References

Cowan, N. (1989a). The reality of cross-modal Stroop effects. *Perception & Psychophysics, 45*, 87-88.

Cowan, N. (1989b). A reply to Miles, Madden, and Jones (1989): Mistakes and other flaws in the challenge to the cross-modal Stroop effect. *Perception & Psychophysics, 45*, 82-84.

Cowan, N., & Barron, A. (1987). Cross-modal, auditory-visual Stroop interference and possible implications for speech memory. *Perception & Psychophysics, 41,* 393-401.

De Jong, R., Liang, C., & Lauber, E. (1994). Conditional and unconditional automaticity: A dual-process model of effects of spatial stimulus-response correspondence. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 731-750.

Dittrich, K., & Stahl, C. (2011). Nonconcurrently presented auditory tones reduce distraction. *Attention, Perception & Psychophysics,73*, 714-719.

Donohue, S. E., Liotti, M., Perez, R. III., & Woldorff, M. G. (2012). Is conflict monitoring supramodal? Spatiotemporal dynamics of cognitive control processes in an auditory Stroop task. *Cognitive, Affective, & Behavioral Neuroscience, 12,* 1-15.

Elliott, E. M., & Cowan, N. (2001). Habituation to auditory distractors in a cross-modal, color word interference task. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 27*, 654-667.

Elliott, E. M., & Cowan, N. (2005). Coherence of the irrelevant-sound effect: Individual profiles of short-term memory and susceptibility to task-irrelevant materials. *Memory & Cognition, 33,* 664-675.

Elliott, E. M., Cowan, N., & Valle-Inclan, F. (1998). The nature of cross-modal, color-word interference effects. *Perception & Psychophysics, 60,* 761-767.

Elliott, E. M., Barrilleaux, K. M. & Cowan, N. (2006). Individual differences in the ability to avoid distracting sounds. *European Journal of Cognitive Psychology, 18*, 90-108.

Green, E. J., & Barber, P. J. (1981). An auditory Stroop effect with judgments of speaker gender. *Perception & Psychophysics, 30,* 459-466.

Hanauer, J. B., & Brooks, P. J. (2003). Developmental change in the cross-modal Stroop effect. *Perception & Psychophysics, 65,* 359-366.

Hughes, R. W., Vachon, F., & Jones, D. M. (2007). Disruption of short-term memory by changing and deviant sounds: Support for a duplex-mechanism account of auditory distraction. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 33*, 1050-1061.

Hutchison, K. A. (2011). The interactive effects of listwide control, item-based control, and working memory capacity on Stroop performance. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 37*, 851-860.

Kahneman, D., & Chajczyk, D. (1983). Tests of the automaticity of reading: Dilution of Stroop effects by color-irrelevant stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 9*, 497-509.

Kane, M. J., & Engle, R. W. (2003). Working memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132,* 47-70.

Lavie, N. (2005). Distracted and confused? Selective attention under load. *Trends in Cognitive Science, 9*, 75-82.

Logan, G. D. & Zbrodoff, N. J. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory & Cognition, 7*, 166-174.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163-203.

Meier, M. E., & Kane, M. J. (2013). Working memory capacity and Stroop interference: Global versus local indices of executive control. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 39,* 748-759. doi:10.1037/a0029200

Miles, C., & Jones, D. M. (1989). The fallacy of the cross-modal Stroop effect: A rejoinder to Cowan (1989). *Perception & Psychophysics, 45*, 85-86.

Miles, C., Madden, C., & Jones, D. M. (1989). Cross-modal, auditory-visual Stroop interference: A reply to Cowan and Barron (1987). *Perception & Psychophysics, 45*, 77-81.

Morey, C. C., Elliott, E. M., Wiggers, J., Eaves, S. D., & Shelton, J. T. (2012). Goal neglect links Stroop interference with working memory capacity. *Acta Psychologica, 141,* 250-260. doi:10.1016/j.actpsy.2012.05.013

Parmentier, F. B. R., Elsley, J. V., Andrés, P., & Barceló, F. (2011). Why are auditory novels distracting? Contrasting the roles of novelty, violation of expectation and stimulus change. *Cognition, 119*, 374–380.

Pratte, M.S., Rouder, J.N., Morey, R.D., & Feng, C. (2010). Exploring the differences in distributional properties between Stroop and Simon effects using delta plots. *Attention, Perception, and Psychophysics, 72*, 2013-2025.

Ridderinkhof, K. R., Scheres, A., Oosterlaan, J., & Sergeant, J. A. (2005). Delta plots in the study of individual differences: New tools reveal response inhibition deficits in AD/HD that are eliminated by methylphenidate treatment. *Journal of Abnormal Psychology, 114*, 197-215.

Risko, E. F., Schmidt, J. R., & Besner, D. (2006). Filling a gap in the semantic gradient: Color associates and response set effects in the Stroop task. *Psychological Bulletin & Review, 13*, 310-315.

Roberts, K. L., & Hall, D. A. (2008). Examining a supramodal network for conflict processing: A

systematic review and novel functional magnetic resonance imaging data for related visual and auditory Stroop tasks. *Journal of Cognitive Neuroscience, 20,* 1063-1078.

Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review, 110*, 88-125.

Roelofs, A. (2005). The visual-auditory color-word Stroop asymmetry and its time course. *Memory & Cognition, 33,* 1325-1336.

Roelofs, A. (2010). Attention and facilitation: Converging information versus inadvertent reading in Stroop task performance. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 36*, 411-422.

Roelofs, A. (2012). Attention, spatial integration, and the tail of response time distributions in Stroop task performance. *Quarterly Journal of Experimental Psychology, 65*, 135-150.

Schmidt, J. R., & Besner, D. (2008). The Stroop effect: Why proportion congruence has nothing to do with congruency and everything to do with contingency. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 34*, 514-523.

Schmidt, J. R., & Cheesman, J. (2005). Dissociating stimulus-stimulus and response-response effects in the Stroop task. *Canadian Journal of Experimental Psychology, 59*, 132-138.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools Inc.

Schröger, E., & Wolff, C. (1998). Behavioral and electrophysiological effects of task-irrelevant sound change: A new distraction paradigm. *Cognitive Brain Research, 7*, 71-87.

Shimada, H. (1990). Effect of auditory presentation of words on color naming: The intermodal Stroop effect. *Perceptual and Motor Skills, 70*, 1155-1161.

Speckman, P., Rouder, J. N., Morey, R. D., & Pratte, M. (2008). Delta plots and coherent distribution ordering. *American Statistician, 62*, 262-266.

Spieler, D. H., Balota, D. A., & Faust, M. E. (2000). Levels of selective attention revealed through analyses of response time distributions. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 506-526.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643-662.

Unsworth, N., Redick, T. S., Spillers, G. J., & Brewer, G. A. (2012). Variation in working memory capacity and cognitive control: Goal maintenance and microadjustments of control. *Quarterly Journal of Experimental Psychology, 65*, 326-355.

Table 1

Percentages of Errors by Distractor Condition in Each of the Four Versions of the Stroop Task.

	Congruent	Incongruent	Silence	NonColor
CrossModal	0.86	3.22	1.28	1.71
CrossModal@	0.74	1.79	1.19	1.64
MultiModal	0.43	9.77	0.62	0.74
Traditional	0.46	15.04	0.35	1.58

Figure Captions

Figure 1. Examples of the targets and distractors in the four versions of the Stroop task. The incongruent distractor condition is depicted here, while the complete design included congruent, non-color word, and silent/neutral distractor conditions as well.

Figure 2. Means of the median reaction times for the four types of distractor conditions in each of the four versions of the Stroop task. Error bars represent standard error of the mean.

Figure 3. Means of the median reaction times for the four types of distractor conditions in the cross modal task, by block. Error bars represent standard error of the mean.

Figure 4. Distributional RT analyses. Panel A compares incongruent and congruent distributions (i.e., Stroop effect) for each task. Panels B, C, and D (respectively) compare incongruent, non-color word, and congruent distributions with neutral trials.

Figure 1

Task	Visual Stimulus	Auditory Stimulus in the Incongruent Condition	Correct Response
Cross-Modal		"red"	"blue"
Cross-Modal @	@@@@	"red"	"blue"
Multi-Modal	red	"red"	"blue"
Traditional	red	N/A	"blue"

Figure 2

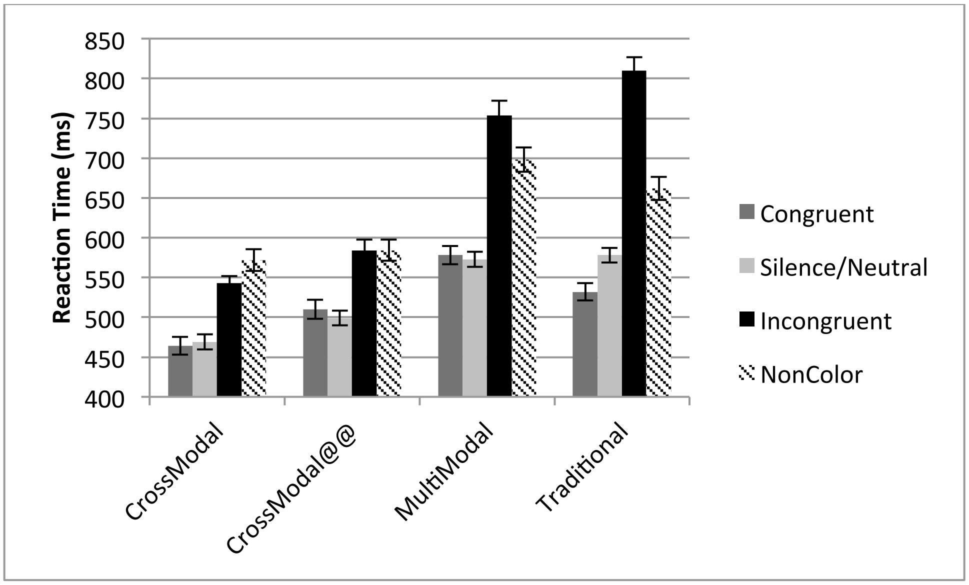

Figure 3